Dual Energy X-ray Material Decomposition

Computer Integrated Surgery II Final Report

May 9, 2019

Cong Gao
Mentors: Mathias Unberath, Mehran Armand, Russell Taylor
the Johns Hopkins University, Baltimore, MD

Contents

1 Introduction **3**
- 1.1 Background . . . 3
- 1.2 Problem . . . 4
- 1.3 Related work . . . 4

2 Methodology **6**
- 2.1 Simulation . . . 6
- 2.2 Architecture . . . 7

3 Experiment **9**
- 3.1 Data acquisition and simulation . . . 9
- 3.2 Pipeline training . . . 10

4 Results **10**

5 Discussion **11**

6 Conclusion **12**

7 Future Work **13**

8 Management Plan **13**

1 Introduction

1.1 Background

Standard X-ray imaging brings difficulty for surgeons to identify region of interest (ROI) features from anatomical clutter. [4] Dual energy X-ray enables anatomical clutter reduction via material decomposition by utilizing the physical properties of X-ray formulations. Fig. 1 illustrates the decomposition process. Traditional Dual Energy X-ray Absorptiometry (DEXA) system has been developed in analyzing bone density, fat tissue, etc. But the model is largely approximated and simplified, because the target ROI is usually large and not targeted to small region accuracy. In the application

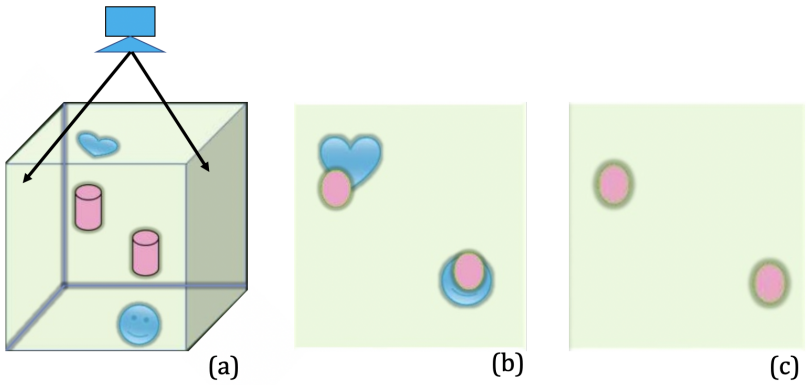

Figure 1: Illustration of material decomposition. (a) 3D objects layout. (b) 2D object clutter. (c) Desired object decomposition.

of femoroplasty operation, surgeons need to rely on X-ray images to identify the injected cement structure. 3D reconstructed femur cement using is fuzzy using cluttered X-rays. A better 2D X-ray decomposition result will help monitor the shape of reconstructed cement during injection. Thus, we propose to re-design the algorithm and improve the accuracy on decomposition of injected cement during femoroplasty. Decomposed frames can be used to improve 3D reconstruction of the injected cement, to better understand the cement distribution.

1.2 Problem

According to Beer-Lambert law, the measurement of X-ray detector can be computed as $I = I^0 e^{-\mu(E)T}$, where I^0 represents the intensity of x-ray photons before entering the materials, E describes the photon energy, and $\mu(E)$ and T represent the attenuation coefficient and the path length. After applying the minus logarithmic operation, the projection m can be expressed as $m = -log(\frac{I}{I^0}) = \mu(E)T$. Considering multiple materials $i = 1...M$, the attenuation integral will then be expressed as $\sum_{i=1}^{M} \mu_i(E)T_i$. In reality, the photon emission is following a spectrum distribution $p_0(E)$, and the attenuation coefficient is dependent on energy E and material density ρ. Thus, we can formulate the above factors in the following expression,

$$m(\mathbf{u}) = \int p_0(E)\Big(\sum_{k\in M} \delta(k, M(\mathbf{x}))(\mu/\rho)_m(E) \int \rho(\mathbf{x}) d\mathbf{l}_\mathbf{u}\Big) dE \tag{1}$$

,where $m(\mathbf{u})$ is the pixel-wise measurement of the X-ray detector and $\delta(k, M(\mathbf{x}))$ is the material identification function in 3D. According to the dependency of energy, Eq. (1) can be further simplified as

$$w_E^M = \int p_0(E)\Big(\sum_{k\in M} \delta(k, M(\mathbf{x}))(\mu/\rho)_m(E)\Big) dE$$

and

$$T(\mathbf{u}) = \int \rho(\mathbf{x}) d\mathbf{l}_\mathbf{u}$$

Thus, we can formulate a linear model of the dual energy measurement $m(\mathbf{u})_L$ and $m(\mathbf{u})_H$ as

$$\begin{cases} m(\mathbf{u})_L = w_L^1 T^1(\mathbf{u}) + w_L^2 T^2(\mathbf{u}) \\ m(\mathbf{u})_H = w_H^1 T^1(\mathbf{u}) + w_H^2 T^2(\mathbf{u}) \end{cases} \tag{2}$$

,where $T^1(\mathbf{u})$ and $T^2(\mathbf{u})$ are the desired pixel-wise decomposition unknowns of two materials. The problem of dual energy X-ray decomposition is to design a model to estimate the unknown $T^i(\mathbf{u}), \{i = 1, ..., M\}$ from the measurements $\{m(\mathbf{u})_L, m(\mathbf{u})_H\}$.

1.3 Related work

Parameters in (2) are calculated based on simulation of spectrum and online published resource of material attenuation. Siemens company has on online

spectrum simulation software https://www.oem-products.siemens-healthineers.com/x-ray-spectra-simulation, where the stop energy of the emitted photon and material of the X-ray machine emission node can be entered as parameters to generate the spectrum. The material attenuation coefficients are referred to the NDI published resource https://www.nde-ed.org/EducationResources/CommunityCollege/Radiography/Physics/attenuationCoef.htm. In our problem, we specifically use bone and iodine, because the injected cement contains iodine element. Thus, w_E^M in (2) can be interpolated between the acquired spectrum and attenuation coefficient.

Based on the model we presented, the naive solution is solving a least square problem for each pixel $\mathbf{u}$:

$$min_{T(\mathbf{u})}\big(M(\mathbf{u}) - WT(\mathbf{u})\big)^T\big(M(\mathbf{u}) - WT(\mathbf{u})\big) \tag{3}$$

, where

$$M = \begin{bmatrix} m(\mathbf{u})_L \\ m(\mathbf{u})_H \end{bmatrix}, T = \begin{bmatrix} T^1(\mathbf{u}) \\ T^2(\mathbf{u}) \end{bmatrix}, W = \begin{bmatrix} w_L^1 & w_L^2 \\ w_H^1 & w_H^2 \end{bmatrix}.$$

The problem of least square solution is overfitting to particular training data. Because the coefficient matrix W is acquired based on modeling, the artifacts will bias the result T to unreasonable solutions, like negative numbers. Thus, we add an non-negative constraint to the least square objective,

$$min_{T(\mathbf{u})}\big(M(\mathbf{u}) - WT(\mathbf{u})\big)^T\big(M(\mathbf{u}) - WT(\mathbf{u})\big) \text{ s.t. } T(\mathbf{u}) \geq \mathbf{0} \tag{4}$$

This constrained least square solution is considered as model-based solution for baseline comparison. We then introduce deep learning to combine with the model-based solution to improve the result.

In a recent study of computed tomography (CT) reconstruction [1], the author presented a novel method to combine model-driven approach and data driven approach for solving ill-posed inverse problems. It gave state-of-the-art results on computed tomography problem for both analytical and human phantoms. The innovation is introducing deep learning to replace the proximal operator in the traditional PDHG algorithm by a network [2]. This design is then called Learned PDHG algorithm. Fig. 2 displays the details of this learned PDHG algorithm. The operator Γ_θ and Λ_θ are called learned proximal, which simply notes the learning part. This idea is inspirational to

our material decomposition pipeline. We can also bring this insight to our design by learning the updates of the variables in primal/dual domain using a convolutional network.

Algorithm 3 Learned Primal-Dual

Initialize $f_0 \in X^{N_{\text{primal}}}, h_0 \in U^{N_{\text{dual}}}$
for $i = 1, \ldots, I$ **do**
  $h_i \leftarrow \Gamma_{\theta_i^d}\left(h_{i-1}, \mathcal{K}(f_{i-1}^{(2)}), g\right)$
  $f_i \leftarrow \Lambda_{\theta_i^p}\left(f_{i-1}, [\partial\mathcal{K}(f_{i-1}^{(1)})]^*(h_i^{(1)})\right)$
return $f_I^{(1)}$

Figure 2: Learned Primal-Dual Algorithm.

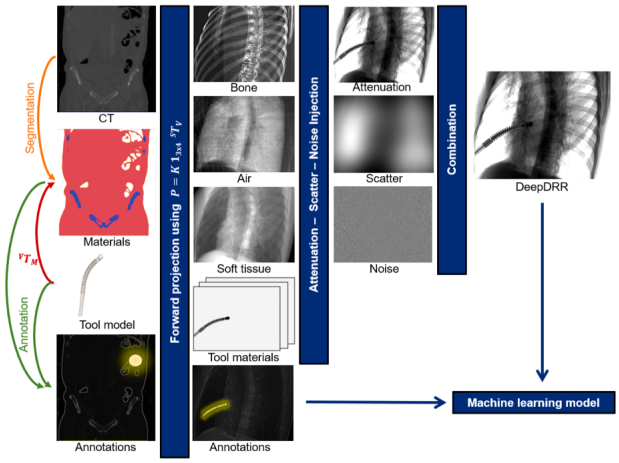

Figure 3: DeepDRR Pipeline [5]

2 Methodology

2.1 Simulation

In order to conduct simulation study, we want to use the recently proposed X-ray simulation framework DeepDRR, which can perform fast and realistic

simulation of fluoroscopy and digital radiography from CT scans by using deep learning [5]. Fig. 3 is the pipeline of DeepDRR. We can also simulate X-ray projections with different energy levels. Another benefit by introducing DeepDRR is that it enables segmentation in 3D domain, including bone, soft tissue and air, which can be used to generate target decomposition projections as groundtruth images for training.

The CT image is acquire from *ex vivo* femurs during our lab cadaver study. We only use the left femur for simulation. We also simulate 3D cement model and manually place the cement in the femur head to do DRR projection. Fig. 4 shows the DRR simulation effect.

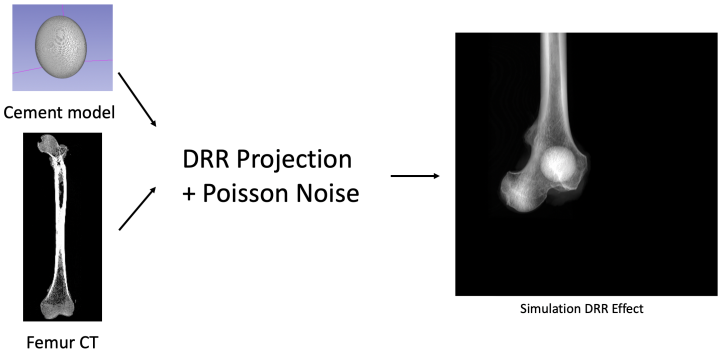

Figure 4: DRR simulation effect.

2.2 Architecture

We formulate the problem as the following optimization objective, where the first term is the regular least square model, which represents the physical solution; $L(T(\mathbf{u}))$ models the non-negative physical constraint of decomposition result; $L_N(T)$ is the learned regularization, which contains spatial information.

$$min_{T(\mathbf{u})}\big(M(\mathbf{u}) - WT(\mathbf{u})\big)^T\big(M(\mathbf{u}) - WT(\mathbf{u})\big) + L(T(\mathbf{u})) + L_N(T) \qquad (5)$$

We then put the above objectives into a single deep network to perform optimization concurrently. Fig. 5 presents the proposed architecture. The original input is X-ray acquisitions "Low" and "High". The Generator is a

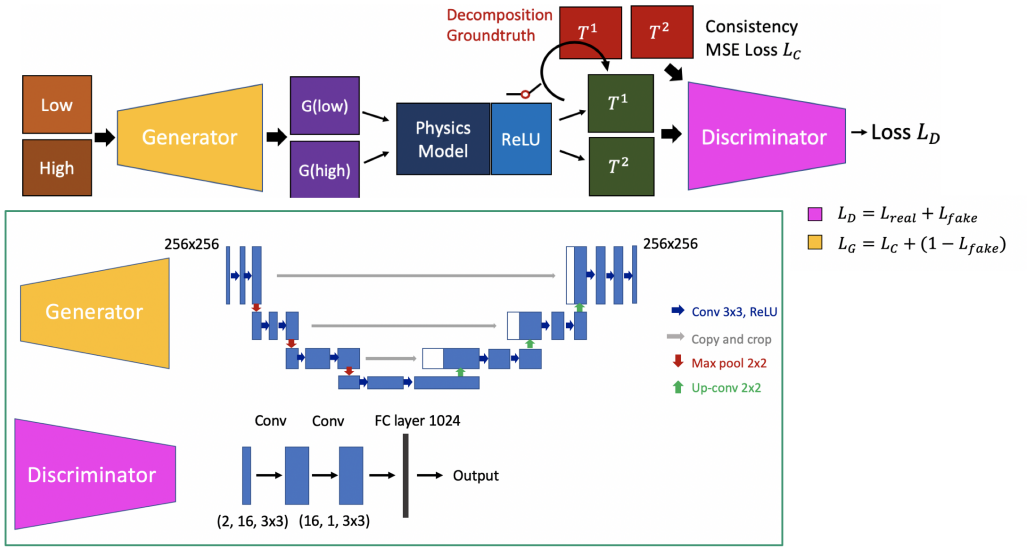

Figure 5: Proposed pipeline of training material decomposition.

U-net [3] based autoencoder, predicting the update of input images. The Discriminator are basically two convolutional layers and two Fully connected layers. Because of the memory limitation, we resize the original image to be 256×256 resolution. The output of Generator is the update images, noted G(low) and G(high). They are then passed through the physics model to convert it into decomposition domain. The ReLU activation is then connected as non-negative constraint. Now, we get the predicted decompositions $\{T^1, T^2\}$. They and their paired groundtruth images are both sent to Discriminator. The loss of Discriminator is calculated as $L_D = L_{real} + L_{fake}$ for each paired input. In order to drive the Generator prediction closer to groundtruth, we add another consistency loss L_c, which is mean squared error of prediction and groundtruth for generator. Thus, the generator loss is composed of $L_G = L_C + (1 - L_{fake})$.

3 Experiment

3.1 Data acquisition and simulation

We conducted cadaver study of femur cement injection using the nView system. It provides low-dose X-ray projections and fast 3D reconstruction, which can later be used for validation purpose. The stop energies are set as 60 kV and 75 kV. For each single acquisition, the nView system automatically acquires 125 projections equally distributed along its base round emitter track. Because of the time sensitivity of the injection process, we acquired dual energy projections twice during and after the injection.
Fig. 6 shows our experiment setup. We use the same system parameters to

Figure 6: Experiment setup of femur cement injection using the nView system.

simulate the X-ray spectrum using the Siemens software as mentioned before. It uses tungsten anode and 3 mm aluminum plus 0.5 mm copper filtration. The stop energies are the same 60 kV and 75 kV. Fig. 7 shows the parameter simulation result. From the figures, we clearly observe the K-edge effect of the iodine element, which is helpful for differentiate the injected cement.

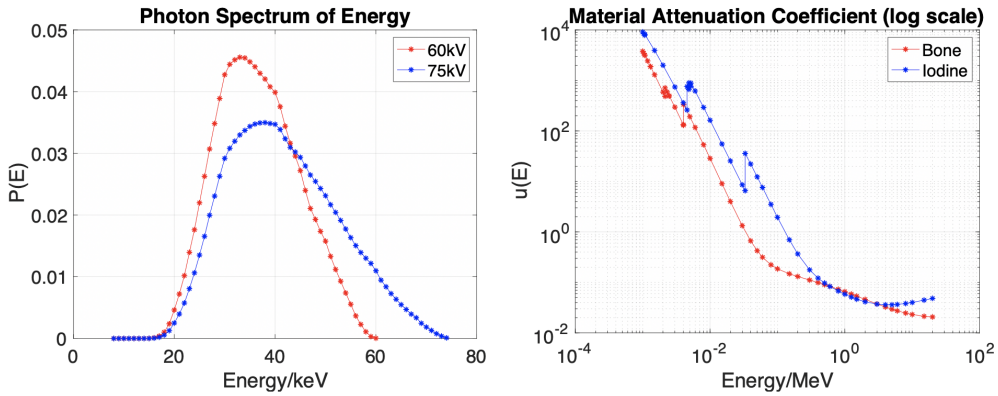

Figure 7: Simulated spectrum and material attenuation coefficients.

3.2 Pipeline training

We simulated in total 2,000 training images and split 1,800 for training and 200 for validation. The simulated X-rays have 976×976 pixels with an isotropic pixel size of $(0.31\text{mm})^2$. Source-to-detector distance was fixed to 800 mm while source-to-isocenter distance was 650 mm. Source rotation in LAO/RAO was $\in [0°, 360°]$ and in CRAN/CUAD $\in [75°, 105°]$. Images are resized to 256×256. Learning rate was set to 0.0001. Batchsize was set to 1. The Generator and Discriminator were updated once per iteration. Training period ran for 100 epochs.

4 Results

We first tested the results on the synthetic validation dataset. The accuracy is reported as mean dice score compared with the groundtruth decomposition and the signal variation. Since the predictions all have background noise, we are using a threshold of 1.2 times mean response to filter the noise. Then we compared the iodine prediction mask with the groundtruth mask. The mean dice score for model based method is 0.268, while our learning based method reaches 0.791. We also calculated the mean standard deviation of the output iodine response. The model based method has mean STD of 0.0037, while our learning based method is 0.0028. So the learning based method is largely outperforming model based method in prediction accuracy and smoothness of response.

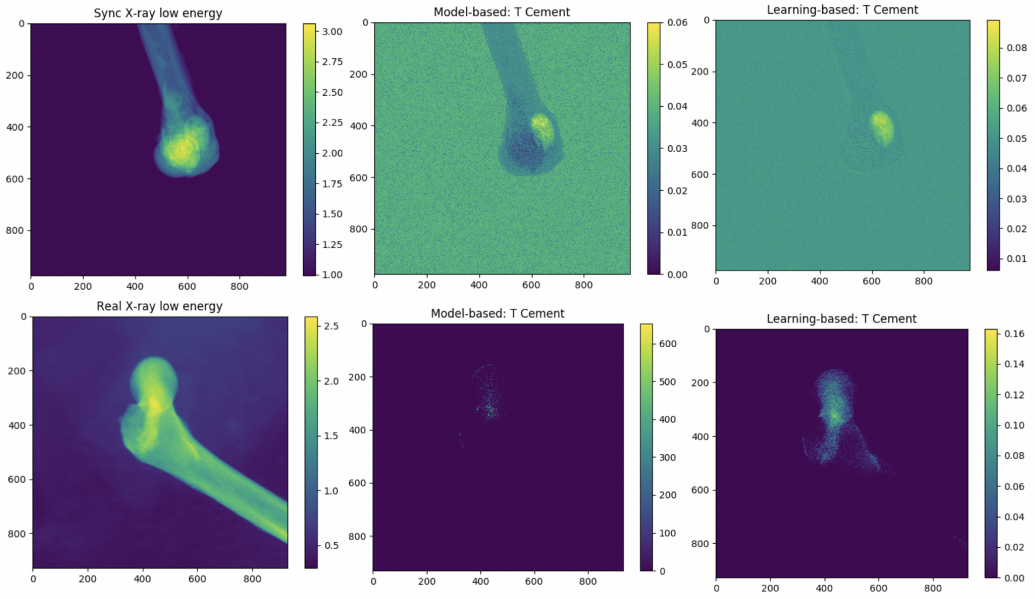

Figure 8: Decomposition results of cement. Upper row: synthetic data; lower row: real data.

We then tested the model on the unseen real X-rays. The constrained least squared solution has very sparse and limited signals, which are mostly because of the overfitting of the linear model. Our learning-based model yields much better result in the cement region. Fig. 8 presents the synthetic validation result and the real X-ray cement decomposition result. We can clearly see that the model based version has stronger noise and bad background prediction in the bone region, while the learning based method has more smooth signals and flatten background, which is closer to the real solution. The real X-ray decomposition also shows the shape of the injection tunnel and the strong distribution at the center of the head, which corresponds the real experiment condition.

5 Discussion

The success of our proposed learning based method over traditional model based method is because 1) DeepDRR simulation is close to the real X-ray effect. It includes the physics factors like photon spectrum, scattering, ma-

terial segmentation, noise generation etc. into consideration, which performs better than conventional DRR simulation environment. With the simulated poisson photon noise and electric noise, it is close to the real X-ray noise distribution, which drives the network to update in the correct direction. 2) We explicitly include the physical model and constraints into our pipeline design and predict exactly the update residual term of the input projections, which largely regulates the network to focus on the critical part of the problem. It will then compensate the non-ideal model simulation and updates the input projections, rather than overfit to the specific training data features. This helps to increase the generalization ability of the model against traditional model based method. 3) The GAN based architecture also improves the generator ability during the adversial iteration. The consistency loss also helps to drive the generator along the correct track, instead of easily converging to a bad local minimum.

Because of time constraint, we didn't investigate much into a more complex physical model. Currently, our baseline model is an approximated linear combination, which is actually an naive model. A better model will then consider energy dependent spectrum and attenuation coefficient relationship. By integrating a more complex model, it will also require improvement of the current pipeline to include the update of the model parameters during training process. The current physical model has fixed the parameters and only updates the network weights. We are also limited in cement model and simulation data variants. A more abundant dataset will also improve the results. The real X-ray stop energies are now set to 60 kV and 75 kV, which are essentially a small gap in the application of dual energy X-ray decomposition. It will also help improve the performance with a larger energy gap.

Overall the result is encouraging, because it inspires the research to introduce learning-based regularization in solving ill-posed inverse problem in mathematical physicals, especially in medical physics. This work has the potential to generalize to other related inverse problem if investigated enough.

6 Conclusion

Look into more complex physical model instead of a linear version. The pipeline was trained on synthetic dataset generated using DeepDRR. Perfor-

mance is evaluated using real X-rays acquired during femur cement injection cadaver study. The cement decomposition result is largely improved compared to traditional model-based methods. This research shows preliminary proof-of-concept results to introduce learning-based regularization in solving ill-posed inverse problem in mathematical physics.

7 Future Work

This work will be extended during summer research and beyond. Possible directions include

- Look into more complex physical model instead of a linear version.
- Include physical model parameters into training and test with more complex pipeline.
- Test more than two material situation. It is essentially still an ill-posed inverse problem. The same insight could be generalized to solve harder multiple material decomposition.
- After the pipeline is robust in segmenting the cement in 2D dual energy Xrays, we will then try to use the decomposition result to reconstruct the 3D shape of the cement and ideally monitor the 3D shape during the injection process.

8 Management Plan

I have a weekly discussion with Dr.Unberath about the design and progress. I also update Dr.Armand and Dr.Taylor in one-on-one meeting or group meeting about milestones. The code is primarily developed on a work desktop independently by Cong Gao. The current source code is now documented and uploaded to a private github account for backup purpose.

I want to express my thanks to graduate student Mahsan Bakhtiari and Amir Farvardin's help in conducting femur injection experiment using the nView system.

References

[1] Adler, J., Öktem, O.: Learned primal-dual reconstruction. IEEE transactions on medical imaging **37**(6), 1322–1332 (2018)

[2] Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of mathematical imaging and vision **40**(1), 120–145 (2011)

[3] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)

[4] S Saija, F Ursani, A.U., Paul, N.S.: Dual-energy imaging and digital tomosynthesis: Innovative x-ray based imaging technologies (2016)

[5] Unberath, M., Zaech, J.N., Lee, S.C., Bier, B., Fotouhi, J., Armand, M., Navab, N.: Deepdrr–a catalyst for machine learning in fluoroscopy-guided procedures. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 98–106. Springer (2018)